water

MDPI

Article

Uncertainty Analysis of the Water Scarcity Footprint Based on the AWARE Model Considering Temporal Variations

Jong Seok Lee [1], Min Hyeok Lee [1], Yoon-Young Chun [2] and Kun Mo Lee [1,*]

[1] Department of Environmental and Safety Engineering, Ajou University, Worldcup-ro 206, Yeongtong-gu, Suwon 16499, Korea; jongseok007@ajou.ac.kr (J.S.L.); wcwsky@ajou.ac.kr (M.H.L.)
[2] Research Institute of Science for Safety and Sustainability, National Institute of Advanced Industrial Science and Technology, 16-1 Onogawa, Tsukuba 305-8569, Japan; yy.chun@aist.go.jp
* Correspondence: kunlee@ajou.ac.kr; Tel.: +82-31-219-2405; Fax: +82-31-219-1613

Received: 20 February 2018; Accepted: 17 March 2018; Published: 19 March 2018

Abstract: The purpose of this paper is to compare the degree of uncertainty of the water scarcity footprint using the Monte Carlo statistical method and block bootstrap method. Using the hydrological data of a water drainage basin in Korea, characterization factors based on the available water remaining (AWARE) model were obtained. The uncertainties of the water scarcity footprint considering temporal variations in paddy rice production in Korea were estimated. The block bootstrap method gave five-times smaller percentage uncertainty values of the model output compared to that of the two different Monte Carlo statistical method scenarios. Incorrect estimation of the probability distribution of the AWARE characterization factor model is what causes the higher uncertainty in the water scarcity footprint value calculated by the Monte Carlo statistical method in this study. This is because AWARE characterization factor values partly follows discrete distribution with extreme value on one side. Therefore, this study suggests that the block bootstrap method is a better choice in analyzing uncertainty compared to the Monte Carlo statistical method when using the AWARE model to quantify the water scarcity footprint.

Keywords: water scarcity footprint; characterization factor; uncertainty analysis; Monte Carlo statistical method; block bootstrap; probability distribution

1. Introduction

Increasing interest in water footprints has led to the publication of the international water footprint standard, ISO 14046 [1]. A water footprint consists of two major pillars: water quality degradation and water consumption [1,2]. From the perspective of a life cycle assessment (LCA), the emission and water consumption of a product and/or organization are called its life cycle inventory. The potential impact of environmental emissions is a function of the background concentration of the affected area. Likewise, the potential impact of water consumption is a function of the water demand and availability in the affected area [3]. In the context of ISO 14046, the water scarcity footprint (WSF) is the potential impact associated with the quantity of water consumption, without considering the water quality [1,4]. Water consumption can be divided into blue (surface and groundwater) and green (precipitation) water from the perspective of water scarcity [5]. However, the impact pathway of green water consumption is not well known [6] and it could double count with land use impact category [7]. Therefore, this study only focused on blue water consumption.

To estimate the potential impact of blue water consumption, the water scarcity characterization factors (CFs) should consider local conditions [1,5]. Most water scarcity CFs take into account

local climatic characteristics such as drainage basins (spatial) and monthly variations (temporal) in water supply and demand. There are many approaches to defining the CF of water scarcity such as withdrawal to availability (WTA), consumption to availability (CTA), and availability minus demand (AMD).

The CFs of WTA [8–13] and CTA [14–18] are based on the ratio of human water withdrawal to hydrological water availability and the ratio of human water consumption to hydrological water availability, respectively. The hydrological water availability of a drainage basin reflects the renewable freshwater volume, which can be quantified using the long-term runoff [14,19]. Water withdrawal is the anthropogenic removal of water from any water body or drainage basin, temporarily [1]. Water consumption is often used to describe the water that was removed from but not returned to the same drainage basin [4,14–18]. Withdrawals can include large volumes of water that are returned to the basin immediately after removal, which leads to an overestimation of the water scarcity [14]. Therefore, it has been argued that a consumption-based (CTA) CF is more relevant than a withdrawal-based (WTA) CF [15,19].

The CF of AMD [4] is based on the difference between the hydrological water availability and water demand for both humans and ecosystem, which is the absolute scarcity (remaining availability per area). This CF is also called available water remaining (AWARE). Recently, the United Nations Environment Program/Society of Environmental Toxicology and Chemistry (UNEP/SETAC) recommended using the AWARE CF for water scarcity [3].

A major drawback of the ratio approach (e.g., WTA or CTA) to quantifying the WSF is its inability to represent the absolute water availability [4,20]. This sometimes leads to arid areas with less scarcity than known water-abundant regions [6,19]. In addition, Boulay et al. [4] stated that a stochastic uncertainty assessment of AWARE should be carried out in the future. Therefore, the AWARE CF model was chosen in this study to quantify the uncertainty of the WSF.

From the perspective of the WSF, uncertainties are mostly associated with spatial and temporal characteristics [1]. However, most WSF studies did not consider the uncertainty of the WSF [8–18]. Pfister and Hellweg, Nunez et al., and Scherer and Pfister estimated the uncertainty of water stress index CF [21,22] and water scarcity CF [23] using the Monte Carlo statistical method (MCS). Pfister and Hellweg [21] and Nunez et al. [22] assumed the parametric probability distribution of the input variables (e.g., availability and consumption) and analyzed the uncertainty of the water stress index CF. Scherer and Pfister [23] assumed a normal distribution for six global models (e.g., WaterGap2) and analyzed the model uncertainty of the water scarcity CF.

Although MCS is the method of choice in uncertainty analysis of the WSF studies [21–23], its use for uncertainty analysis has limitations. It requires defining the probability distribution of each input variable, which might be more difficult if empirical information is unavailable [24]. Groen et al. [25] compared four different uncertainty analysis methods (Monte Carlo sampling, Latin hypercube sampling, Quasi-Monte Carlo sampling, and analytical uncertainty propagation) for greenhouse gas emissions and found that the relative uncertainty results of the four methods are not significantly different. However, they assumed that the parametric probability distributions are either normal or lognormal and stated that these assumptions were the limitations of their work. Based on Chen and Corson [24], expert judgment can help determine the statistical parameters of the distributions if there is a lack of empirical information; however, this represents an additional source of uncertainty. Therefore, other uncertainty methods should be considered for analyzing the uncertainty of the WSF if the probability distribution of the input data is uncertain.

The objective of this research is to compare the uncertainty of the WSF associated with temporal variability using two different uncertainty analysis methods: MCS and block bootstrap. Paddy rice production in Korea was used as the case study for WSF uncertainty analysis in this study. A salient feature of this study is that it is the first to compare the bootstrap method with the MCS method to estimate uncertainty of the WSF.

The Materials and Methods section describes the data collection for AWARE CFs and irrigation water consumption for the paddy rice production in a specific river basin in Korea. Moreover, it describes the application of two uncertainty methods: MCS and block bootstrap. The Results and Discussion section details the results of using the two uncertainty analysis methods to calculate the WSF of the paddy rice production. The Conclusion section explains the contribution of this research to the field and future research directions.

2. Materials and Methods

2.1. Available Water Remaining (AWARE) Water Scarcity CF Model

The AWARE model quantifies the potential of water deprivation of either humans or ecosystems and is used to calculate the WSF based on ISO 14046. Equations (1) and (2) show the definition of AMD and CF, respectively [3,4]. The model is based on the available water remaining per unit of surface in a given watershed relative to the world average after human and aquatic ecosystem demands have been met.

$$\mathrm{AMD}_{i,j} = \frac{\mathrm{A}_{i,j} - \left(\mathrm{C}_{i,j} + \mathrm{EWR}_{i,j}\right)}{\mathrm{Area}_j} \quad (1)$$

$$\mathrm{CF}_{i,j} = \frac{1/\mathrm{AMD}_{i,j}}{1/\mathrm{AMD}_{\text{world avg}}} = \frac{\mathrm{AMD}_{\text{world avg}}}{\mathrm{AMD}_{i,j}} \quad (2)$$

Here, $\mathrm{A}_{i,j}$ is the water availability in the ith month in region j (m^3/month), $\mathrm{C}_{i,j}$ is the human water consumption (m^3/month), $\mathrm{EWR}_{i,j}$ is the environmental water requirement (m^3/month), Area_j is the jth area (m^2), and $\mathrm{AMD}_{\text{world avg}}$ is the world average AMD ($m^3/(m^2 \cdot \text{month})$).

The sum of human water consumption and environmental water requirement is referred to as demand. The unit of $\mathrm{AMD}_{i,j}$ is $m^3/(m^2 \cdot \text{month})$; $\mathrm{CF}_{i,j}$ is dimensionless, expressed as $m^3_{\text{world-eq}}/m^3_{i,j}$. When $\mathrm{AMD}_{i,j}$ is 100 times less than the AMD world average, the CF is given the maximum value of 100. When $\mathrm{AMD}_{i,j}$ is 10 times greater than the AMD world average, CF is given the minimum value of 0.1 [4].

2.2. Case Study for Uncertainty Analysis

As defined earlier, the WSF is the potential impact associated with the quantity of the water consumption. The functional unit (fu) of paddy rice production was 1 kg of rice. The WSF of the paddy rice production per functional unit was calculated using the linear WSF model shown in Equation (3).

$$\mathrm{WSF} = \mathrm{X}\mathrm{g} \quad (3)$$

Here, WSF ($m^3_{\text{world-eq}}$/fu) is the potential impact of the water consumption by paddy rice production per functional unit, X is the vector of the mean monthly CF_i ($m^3_{\text{world-eq}}/m^3_{\text{region}}$, 1×12) based on the AWARE model, and g is the inventory vector (m^3_{region}/fu, 12×1) representing the monthly load of the paddy rice production.

Because the focus of this paper is to assess the uncertainty of the WSF, Geum River Basin in Korea was chosen for the case study. It is assumed that the water consumption for the paddy rice production in a given month is uniform in the chosen basin not considering temporal (year-to-year) variation. Therefore, monthly water consumption of the rice production was treated as constant in the WSF model. However, the CF was treated as a random variable in Equation (3) in order to consider temporal (year-to-year) variation.

2.2.1. Paddy Rice Production

Paddy rice production consumes more than 95% of all water during irrigation [26,27]. The electricity consumption, fuel consumption, and pesticide and fertilizer consumption were therefore excluded from the system boundary of the rice production as shown in Figure 1.

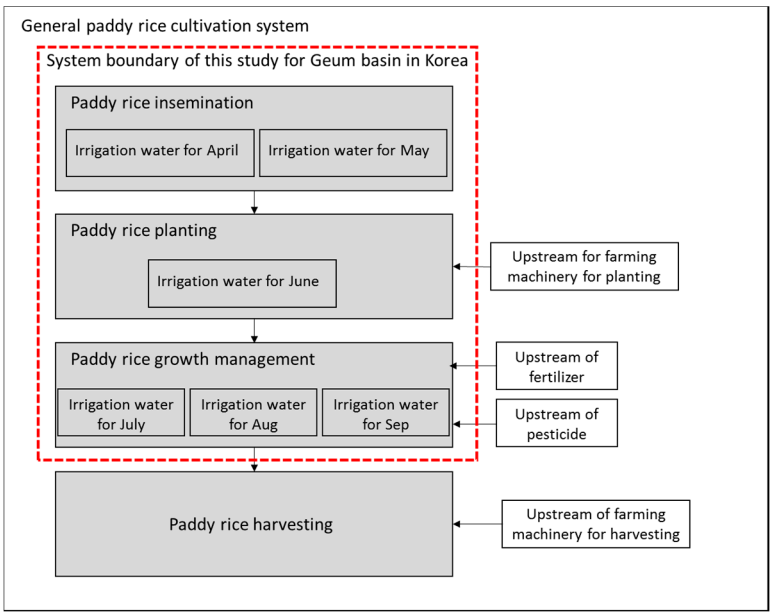

Figure 1. System boundary for the water scarcity footprint of paddy rice production.

The water consumption of rice production was estimated using the annual consumption of irrigation water [28] and annual rice production in the Geum River Basin. Water for rice production is only consumed during irrigation periods (April to September); thus, monthly variations exist. The monthly ratio of irrigation water consumption for irrigation periods for the Geum River Basin was collected from the Rural Agricultural Water Resource Information System [29]. The yearly average water consumption of the rice production was 0.51 m^3_{Geum}/fu.

2.2.2. Characterization Factor Data Sources

The AWARE method uses the WaterGAP2.2 model to estimate the hydrological data (water availability and human demand); data is available for the 34 largest watersheds in the world [30]. However, Nunez et al. [22] encourage updating the water stress index CF using information at the national level because the global model (e.g., WaterGAP2.2) may not represent the local condition of specific basins in a country. Thus, data on the natural runoff and water withdrawal was collected from the Korea National Water Resources Management Systems [28,29]. Based on the collected data, water availability, human demand, and ecosystem demand were estimated.

The water availability was estimated to be the actual runoff, including the human impact on the flow regulation [4]. However, the Korea National Water Resources Management System, which is based on Precipitation-Runoff Modeling System model, only provides natural runoff data, proven to closely mimic the natural water availability in Korea [31]. Therefore, the ratio between actual runoff and natural runoff provided by Water Use in Life Cycle assessment (WULCA, Basin ID 36565) was multiplied with long-term monthly natural runoff data from the Precipitation-Runoff Modeling System and used as the water availability in the Geum River Basin in Korea. The temporal boundary of the water availability was 45 years (1970–2014).

Monthly water withdrawal data for the year of 2014 was used to estimate the human demand to meet the guideline of the LCA initiative, which requires using the water consumption of the current

state [3]. Water withdrawal data was collected from the Water Resource Management Information System [28] in the Geum River Basin. Domestic and industrial water withdrawal and agricultural water withdrawal were converted into water consumption by applying return flow rates of 65% and 35%, respectively [32,33].

Pastor et al. [34] stated that the variable monthly flow method is a valid and easy method for the implementation of global hydrological models. Therefore, in this paper, the environmental water requirement (EWR) was estimated using pristine flow (i.e., natural runoff) with the percentage of variable monthly flow method. It allocates 30–60% of mean monthly flow as a function of seasonal flow pattern: 30% of mean monthly flow to EWR during high-flow seasons when mean monthly flow is over 80% of the mean annual flow, 45% of mean monthly flow during intermediate seasons when mean monthly flow is over 40% and below 80% of the mean annual flow, and 60% of mean monthly flow during low-flow seasons when mean monthly flow is below 40% of mean annual flow [34].

2.3. Uncertainty of the WSF Associated with Temporal Variability

Two different methods were used to estimate the uncertainty associated with temporal variability of the value of the WSF in Equation (3): MCS and block bootstrap. Both methods were used to estimate the variance and mean of the WSF with a confidence interval.

2.3.1. Monte Carlo Statistical Method

The MCS method, a stochastic method to estimate the uncertainty of the model output, calculates the model output based on generated input data. The procedure is repeated many times (e.g., 10,000 iterations) to obtain the probability distribution of the WSF model output from which the mean, variance, and confidence interval of the WSF can be computed. In this study, 10,000 iterations were used according to the recommendations by Martorell et al. [35].

In principle, MCS is performed in four steps [36]: (i) choose the probability distribution of a chosen input variable, X_i, and identify the value of statistical parameters such as the mean and variance if the probability distribution is normal. If it is not, find the relevant probability distribution and values of its statistic; (ii) generate random deviates from a given probability distribution of X_i using the transformation method, where the method transforms the uniform variate of r [0,1] of the uniform distribution to a random deviate X_i of a certain interval of a predetermined probability distribution. This transformation is based on the conservation of the probability of the two distributions; (iii) repeat the previous two steps for all the chosen X_i and compute the model output z for the generated random deviates of all input variables; and finally, (iv) identify the mean ($\overline{Z}$), standard deviation, and confidence interval of the model output at the 95% confidence level. The Spearman coefficient was applied to account for the dependence between monthly CFs. This is because the Spearman correlation measure is robust for non-Gaussian marginal densities [37]. In addition, the correlation coefficient between the input variables is not zero, which indicates that there is some degree of association between the monthly CFs, that is, a nonzero correlation coefficient implies two input variables share some degree of correlation [38].

To apply the MCS method to Equation (3), the equation should be expressed as

$$\text{WSF} = \sum_{i=1}^{12} g_i X_i \tag{4}$$

The X in both Equations (3) and (4) represent CF. However, X in Equation (3) is the mean value of the monthly CFs, while X in Equation (4) represents random variable with annual variation.

Three different scenarios were chosen to compare the effect of different probability distributions of X_i (i.e., CF_i) on the WSF: (i) all input variables follow parametric probability distributions; (ii) all input variables follow nonparametric probability distributions (i.e., empirical distributions);

and (iii) the combined case, where input variables either follow nonparametric or parametric probability distributions.

In the case of the parametric probability distribution, the relevant probability distribution of X_i (i.e., CF_i) in the Geum basin was estimated using the Anderson–Darling method [39]; the parameter values of X_i were estimated. The hypotheses to test the parametric probability distribution using the Anderson–Darling test were that in H_0, the input data follows a specific distribution; and in H_1, the input data does not follow a specific distribution. For the goodness-of-fit test, 13 parametric probability distributions were used: lognormal, Pareto, Weibull, gamma, logistic, extreme value, Student's-*t*, normal, Beta, BetaPERT, exponential, uniform, and triangular. If the *p*-value is less than 0.05, the null hypothesis is rejected. In general, if one cannot estimate the probability distribution, either normal or lognormal distributions are assumed in the LCA [40,41].

The empirical distribution based on the histogram was used to estimate the probability distribution in the nonparametric probability distribution case. For the combined case, the input variables not following a specific distribution were estimated using the empirical distribution.

The value of the WSF, which is the mean of the simulated runs, and variance of the WSF are obtained by solving Equation (4). The output of this method includes the mean ($\overline{\text{WSF}}$), variance, and confidence interval of the WSF. The confidence interval of the model output of the MCS was obtained using the percentile approach, where the upper and lower bounds of 95% confidence interval were the 0.025 and 0.975 quantiles of the model outputs, respectively.

2.3.2. Block Bootstrap Method

The bootstrap method is based on a resampling of the sample dataset with replacement. It calculates the difference between the resampled mean and sample mean repeatedly (e.g., number of iterations R = 1000). In this study, 1000 iterations were used according to the recommendations by Pattengale et al. [42]. The premise of the method is that the difference between the resampled data mean and sample mean is approximately equal to the difference between the sample mean and population mean [43].

The bootstrap method described above or the empirical bootstrap method cannot be applied to dependent data such as in this study because the empirical bootstrap method does not consider covariance. Lahiri [44] stated that independent and identically distributed bootstrap methods fail when it comes to dependent data. It would be a mistake to resample sequential scalar quantities because the reshuffled resamples would break the temporal dependence. Thus, we decided to use the block bootstrap method instead.

The block bootstrap method employed in this study divides the original water scarcity $CF_{i,j}$ data (45 $\times$ 12 matrix) into 45 (years) non-overlapping blocks with a length of 12 (months). This is intended to keep the 12-month data from a given year in a block. The concept of the block bootstrap method is that if a block is long enough (45 years in this study), the dependence of the original data of each monthly CF (water scarcity $CF_{i,j}$) will be preserved in the resampled $CF_{i,j}$ [45]. The block bootstrap method creates a block using paired monthly CF values, subsequently generating replicates through resampling. The sign of the covariance of the original sample and the replicates are the same. Furthermore, the sign of the correlation coefficients of the original sample are the same as those of the resampled replicates. As such, dependence between monthly CF values remains intact. If we resample once (i.e., R = 1), newly resampled $CF_{i,j}$ data is obtained, from which WSF_1^* can be determined. By repeating the same procedure many times (R = 1000), we can obtain 1000 WSF^* values. A 95% confidence interval of the WSF using the block bootstrap is calculated with the basic percentile method, as shown in Equations (5) and (6).

$$\text{CI}_{\text{block bootstrap}} = [\text{WSF} - \delta^*_{0.025}, \text{WSF} - \delta^*_{0.975}] \quad (5)$$

$$\delta^* = \text{WSF}^* - \text{WSF} \quad (6)$$

Here, WSF is the original WSF model output, WSF* is the block bootstrap-resampled WSF model output (1000 samples), and δ^* is the difference between WSF* and the WSF [43]. We compute $\delta^* = \text{WSF}^* - \text{WSF}$ for each block bootstrap-resampled $CF_{i,j}$ and sort them from smallest to biggest. Because $\delta^*_{0.025}$ is at the 97.5th percentile, it corresponds to the 975th element in the δ^* list. Likewise, $\delta^*_{0.975}$ is at the 2.5th percentile, corresponding to the 25th element in the δ^* list.

The statistical programing language *R* was used for the block bootstrapping (the *R* code is provided in the Supplementary Information section). Crystal Ball [46] was used to solve Equation (4) based on MCS.

2.3.3. Comparison of the Uncertainty

To compare the uncertainty obtained from different uncertainty analysis methods, the ratio between the half-width of the 95% confidence interval and the mean of the model output, WSF, was used [40,47]. The ratio, termed U, represents the percentage uncertainty of the model output and is defined as

$$\text{U} = \frac{\text{Half width of the 95\% confidence interval of the model output}}{\text{The mean of the model output}} \times 100(\%) \quad (7)$$

A smaller U value reflects a more precise estimate of the mean of the model output. In other words, a smaller U value indicates less uncertainty of the WSF.

3. Results and Discussion

Table 1 shows the monthly water consumption (load) data of the paddy rice production. Water consumption for paddy rice production only occurred during the irrigation season from April to September; the consumption differs from month to month.

Table 1. The monthly load data for the paddy rice case study in the Geum River Basin (unit: m^3_{Geum}/fu).

	January	February	March	April	May	June	July	August	September	October	November	December
Rice	-	-	-	0.01	0.03	0.17	0.10	0.13	0.07	-	-	-

The CF values of the Geum River Basin calculated using the previously discussed method discussed above are given in Table S1 of the Supplementary information section. The monthly CF data in Table S1 is arranged in a matrix of 45 (years) × 12 (months).

To perform uncertainty analysis with MCS, the parametric probability distribution of monthly CF was estimated by the Anderson–Darling goodness-of-fit test, as stated before. Table 2 shows the probability distribution of the CFs under three different scenarios for the MCS. In case of monthly CFs of six months (January, February, April, June, September, and December), the null hypothesis was rejected for all the tested kinds of parametric probability distributions (H_0 = the input data follows a specific distribution). For the combined (parametric and nonparametric) distribution scenario, the distribution rejected in the null hypothesis test was replaced by the empirical distribution.

Table 2. Probability distribution of the CFs under three different scenarios of the MCS.

Month	Nonparametric	Parametric	Combined
January	Empirical	- *	Empirical
February	Empirical	-	Empirical
March	Empirical	Weibull	Weibull
April	Empirical	-	Empirical
May	Empirical	Weibull	Weibull
June	Empirical	-	Empirical

Table 2. *Cont.*

Month	Nonparametric	Parametric	Combined
July	Empirical	Weibull	Weibull
August	Empirical	Lognormal	Lognormal
September	Empirical	-	Empirical
October	Empirical	Weibull	Weibull
November	Empirical	Weibull	Weibull
December	Empirical	-	Empirical

* Cannot be determined; the *p*-value is less than 0.05 for the null hypothesis test for determining the parametric probability distribution.

Therefore, the uncertainties of the WSF of the paddy rice production were estimated using two different uncertainty analysis methods: MCS with two scenarios (nonparametric and combined) and block bootstrap. Table 3 shows the mean, confidence interval, confidence interval width, and U for the MCS and block bootstrap methods for the WSF of the rice production. As pointed out earlier, two different scenarios for the probability distribution of the input variables were created in the MCS case such that there are two different uncertainty analysis results.

Table 3 shows the statistical values of the WSF for the rice production. The WSF mean values for the nonparametric, combined MCS methods and block bootstrap were similar. The U value, however, is lowest in the case of the block bootstrap method, approximately five times smaller than that of the other two MCS scenarios.

Table 3. The statistical values of the water scarcity footprint for the rice production (unit: $m^3_{world\text{-}eq}/fu$).

Statistic	MCS Scenarios		Block Bootstrap
	Nonparametric	Combined	
$\overline{WSF}$	6.6	6.1	6.0
Confidence interval	[0.9, 25.4]	[0.6, 22.4]	[3.5, 8.1]
Confidence interval width	24.5	21.8	4.6
U (%)	185.3	177.3	38.6

The percentage uncertainty (U) used in this paper is a parameter for estimating the uncertainty of the model output. It is similar to the coefficient of variation (CV); however, the U value is based on the confidence interval, and CV is based on the standard deviation of the model output. The confidence interval focuses on the uncertainty, while the standard deviation is a measure of data dispersion or square root of variance. Thus, internationally recognized institutions such as IPCC use U instead of CV in estimating the uncertainty of the model output. The U value is the only parameter used in this study to evaluate the degree of relative uncertainty among the three different statistical simulation methods. As such, comparing the appropriateness of the three different simulation methods using the value of U is warranted.

The three simulation methods have different logic for generating the values of the input variables; however, they use the same method for estimating the confidence interval of their model outputs. The only difference is the way of generating the values of the input variables. Since the values of the input variables differ, the model output values also differ. The combined MCS identifies the probability distribution of the input variables (parametric and non-parametric), and the distribution thus identified is used to generate the values of the input variables. The generated values are then fed to the model and generate a model output. By repeating the procedure many times (e.g., 10,000 iterations), enough data points are available, forming a pseudo population. The same procedure applies to the non-parametric MCS method, with one exception that the probability distribution is non-parametric. The bootstrap method generates the values of the input variables using the original sample by following the bootstrap

technique. Thus, there is no need to estimate the probability distribution of the input variables which is the essence of the bootstrap method.

The likely reason for the higher U values of the MCS scenarios is the nonparametric nature of the CF distribution. Particularly, the AWARE CF, which has been used for the WSF in this study, may have a nonparametric nature due to the discrete steps in its function [3]. Therefore, the estimated probability distribution for the MCS scenarios of the CF might be incorrect. For the nonparametric probability distribution, the estimation method was based on the histogram approach. This approach also has a drawback in the arbitrary nature of choosing the bin size [48]. According to Chen and Corson [24], expert judgment can help determine the statistical parameters of the distributions if there is a lack of empirical information; however, this represents an additional source of uncertainty, as shown by the higher U values obtained with the MCS scenarios compared with that of the block bootstrap method.

In most LCA studies, the uncertainty was analyzed using the MCS method with parametric probability distribution [49]. In the case of the uncertainty analysis of the WSF, the MCS method has been widely used. However, the parametric probability distribution of the input variables, such as CF, is difficult to estimate and can be incorrect, partly because the CF data can be nonparametric in nature. According to the Jolliet et al. [41], lognormal distributions are applicable to the LCA due to the positive value of CF, and other parameters.

Figure 2 shows the histogram in blue and estimated parametric probability distribution in red for the CF for June and August in the Geum River Basin. The probability distribution in August was estimated to be the lognormal distribution based on the Anderson–Darling test, where the null hypothesis was assumed as a lognormal distribution. The probability distribution in June, however, could not be estimated using the Anderson–Darling test because all null hypotheses were rejected. In this case, the lognormal distribution was assumed, as suggested by the IPCC guideline [40] and Jolliet et al. [41].

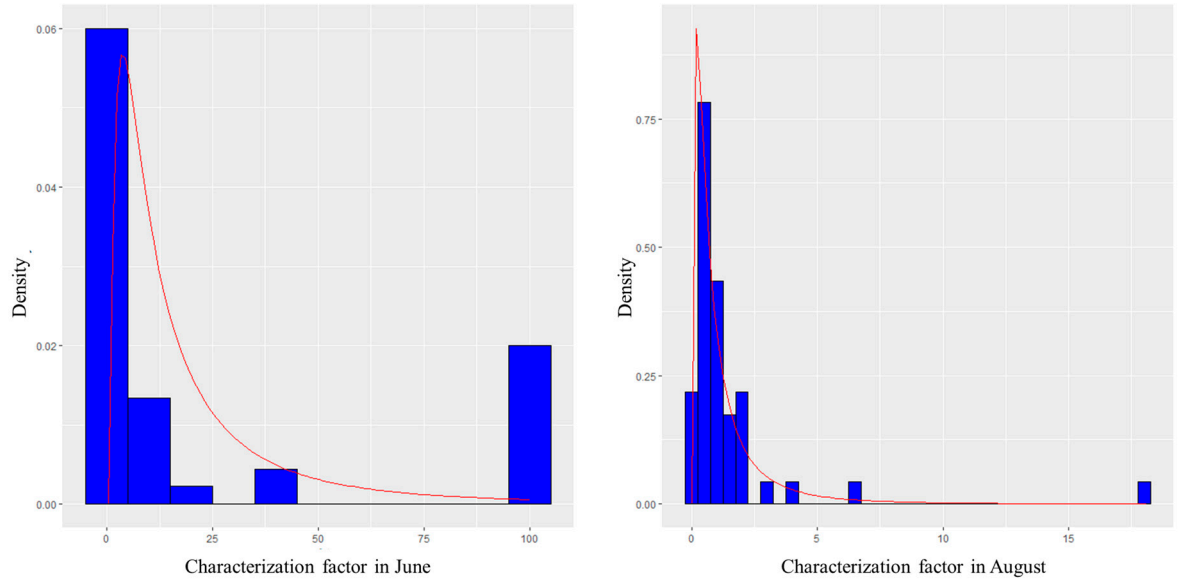

Figure 2. Histogram and adopted lognormal distributions in June and August.

Based on the recommendations of the LCA community [40,41], a lognormal distribution was applied for the "cannot be determined" cases and then compared with the histogram of the CF data. The probability distribution for August in Figure 2 resembles the histogram rather closely. Thus, the estimated lognormal distribution in August closely represents the CF data. However, this is not the case for June. A huge discrepancy can be observed between the estimated lognormal distribution and histogram. The tail side of the data could not capture the peak on the 100 side of the lognormal distribution. In the histogram, there are two peaks on both sides of the X axis and

none in between. This is a typical example of a discrete distribution, where $P(Y_n = y) = 1 - 1/n$ for $y = 0$ and $P(Y_n = y) = 1/n$ for $y = n$ and 0 elsewhere ($P(Y_n = y)$ = probability P at the random variable Y of which value is y, n = integer value) [50]. This indicates that the AWARE CF values partly follow a discrete distribution. Thus, it is a futile attempt to find the parametric distribution of the AWARE CF data where the maximum value is 100. This is because the estimation of the parametric probability distribution may not capture extreme values [47] such as a maximum AWARE CF value of 100. There was no maximum CF value in August; however, in the case of June, 20% (9 years out of 45 years) of the CF data shows a maximum value of 100 (Table S1). Therefore, the parametric probability distribution should not be used when estimating WSF with the maximum value of 100 in the AWARE CF model.

Based on the U values in Table 3 and the discussions above, it can be concluded that the block bootstrap method gave a lower uncertainty than the MCS method in this study. In other words, the block bootstrap method provides a more accurate estimation of the WSF compared with the MCS method, which requires estimation of the parametric and/or nonparametric probability distribution. The block bootstrap method used in this research is a distribution-free method; it uses random sampling of the original CF sample such that there is no step for identifying probability distribution.

Therefore, this study suggests that the block bootstrap method is a better choice in analyzing uncertainty compared to the MCS method in the case of estimating WSF using the AWARE CF model.

4. Conclusions

Uncertainty in the WSF may arise from the temporal variability of the characterization factors. The paddy rice production in the Geum River Basin was used as the case study for WSF uncertainty analysis with MCS (two scenarios) and block bootstrap methods. Percentage uncertainty (U) was lowest in the case of the block bootstrap method, approximately five times smaller than that of the other two MCS scenarios. Estimation of the CF probability distribution is difficult, partly because the CF data can be nonparametric in nature. The MCS method might therefore provide less accurate results due to the incorrect estimation of the probability distribution. This is because AWARE CF values partly follow discrete distribution with extreme value on one side. Therefore, this study suggests that the block bootstrap method is a better choice in analyzing uncertainty compared to the MCS method in the case of estimating WSF using the AWARE CF model. Moreover, the parametric probability distribution should not be used when estimating WSF with a maximum value of 100 in the AWARE CF model.

It is noteworthy that this paper has performed uncertainty analysis of the WSF by considering temporal variability with the block bootstrap method for the first time to estimate the uncertainty of the WSF.

According to ISO 14046, uncertainties are mostly associated with spatial and temporal variations. However, this study focused only on the temporal variability of the monthly CFs. Therefore, uncertainty analysis considering spatial variability with the block bootstrap method will be the future direction of this research. Moreover, this study only treated CF as the random variable. However, in reality, the load (water consumption) should also be treated as a random variable in future studies.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4441/10/3/341/s1.

Acknowledgments: This study was conducted as part of the "Development of the compatible water footprint method" project by Ministry of Trade, Industry and Energy in the Republic of Korea. We express sincere thanks to Professor Jae Eung Yi (Department of Civil Systems Engineering, of the Ajou University) for many valuable discussions.

Author Contributions: Jong Seok Lee conducted research and calculated the water scarcity characterization factor in the Geum River Basin in Korea. Min Hyeok Lee and Yoon-Young Chun performed the *R* programming of the uncertainty analysis methods. Kun Mo Lee conceived the research idea, provided details of the uncertainty analysis methods, and supervised the entire research work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. International Organization for Standardization(ISO). *ISO14046:2014.—Environmental Management—Water Footprint—Principles, Requirements and Guidelines*; ISO: Saint-Denis, France, 2014.
2. International Organization for Standardization (ISO). *ISO/TR 14073:2017—Environmental Management—Water Footprint-Illustrative Examples on How to Apply ISO 14046*; ISO: Saint-Denis, France, 2017.
3. UNEP-SETAC. *Global Guidance for Life Cycle Impact Assessment Indicators Volume 1, Chapter 5a: Water Scarcity*; UNEP/SETAC Life Cycle Initiative: Paris, France, 2016; pp. 100–115, ISBN 978-92-807-3630-4.
4. Boulay, A.-M.; Bare, J.; Benini, L.; Berger, M.; Lathuillière, M.J.; Manzardo, A.; Margni, M.; Motoshita, M.; Núñez, M.; Pastor, A.V.; et al. The WULCA consensus characterization model for water scarcity footprints: Assessing impacts of water consumption based on available water remaining (AWARE). *Int. J. LCA* **2017**. [CrossRef]
5. Quinteiro, P.; Ridoutt, B.G.; Arroja, L.; Dias, A.C. Identification of methodological challenges remaining in the assessment of a water scarcity footprint: A review. *Int. J. LCA* **2017**, *23*, 164–180. [CrossRef]
6. Boulay, A.-M.; Bare, J.; De Camillis, C.; Döll, P.; Gassert, F.; Gerten, D.; Humbert, S.; Inaba, A.; Itsubo, N.; Lemoine, Y.; et al. Consensus building on the development of a stress-based indicator for lca-based impact assessment of water consumption: Outcome of the expert workshops. *Int. J. LCA* **2015**, *20*, 577–583. [CrossRef]
7. Pfister, S.; Boulay, A.-M.; Berger, M.; Hadjikakou, M.; Motoshita, M.; Hess, T.; Ridoutt, B.; Weinzettel, J.; Scherer, L.; Döll, P.; et al. Understanding the LCA and ISO water footprint: A response to hoekstra (2016) "a critique on the water-scarcity weighted water footprint in LCA". *Ecol. Indic.* **2017**, *72*, 352–359. [CrossRef]
8. Frischknecht, R.; Knöpfel, S.B. *Swiss Eco-Factors 2013 According to the Ecological Scarcity Method. Methodological Fundamentals and Their Application in Switzerland*; Federal Office for the Environment, FOEN: Bern, Switzerland, 2013.
9. Frischknecht, R.; Steiner, R.; Jungbluth, N. *The Ecological Scarcity Method Eco-Factors 2006*; Federal Office for the Environment, FOEN: Bern, Switzerland, 2009.
10. Loubet, P.; Roux, P.; Nunez, M.; Belaud, G.; Bellon-Maurel, V. Assessing water deprivation at the sub-river basin scale in lca integrating downstream cascade effects. *Environ. Sci. Technol.* **2013**, *47*, 14242–14249. [CrossRef] [PubMed]
11. Milà i Canals, L.; Chenoweth, J.; Chapagain, A.; Orr, S.; Antón, A.; Clift, R. Assessing freshwater use impacts in LCA: Part I—Inventory modelling and characterisation factors for the main impact pathways. *Int. J. LCA* **2008**, *14*, 28–42. [CrossRef]
12. Pfister, S.; Koehler, A.; Hellweg, S. Assessing the environmental impacts of freshwater consumption in LCA. *Environ. Sci. Technol.* **2009**, *43*, 4098–4104. [CrossRef] [PubMed]
13. Ridoutt, B.G.; Pfister, S. A new water footprint calculation method integrating consumptive and degradative water use into a single stand-alone weighted indicator. *Int. J. LCA* **2013**, *18*, 204–207. [CrossRef]
14. Berger, M.; van der Ent, R.; Eisner, S.; Bach, V.; Finkbeiner, M. Water accounting and vulnerability evaluation (wave): Considering atmospheric evaporation recycling and the risk of freshwater depletion in water footprinting. *Environ. Sci. Technol.* **2014**, *48*, 4521–4528. [CrossRef] [PubMed]
15. Boulay, A.-M.; Bulle, C.; Bayart, J.B.; Deschenes, L.; Margni, M. Regional characterization of freshwater use in LCA: Modeling direct impacts on human health. *Environ. Sci. Technol.* **2011**, *45*, 8948–8957. [CrossRef] [PubMed]
16. Hoekstra, A.Y.; Mekonnen, M.M. *Global Water Scarcity: Monthly Blue Water Footprint Compared to Blue Water Availability for the World's Major River Basins*; Value of Water Research Report Series No. 53; UNESCO-IHE: Delft, The Netherlands, 2011.
17. Hoekstra, A.Y.; Mekonnen, M.M.; Chapagain, A.K.; Mathews, R.E.; Richter, B.D. Global monthly water scarcity: Blue water footprints versus blue water availability. *PLoS ONE* **2012**, *7*, e32688. [CrossRef] [PubMed]
18. Pfister, S.; Bayer, P. Monthly water stress: Spatially and temporally explicit consumptive water footprint of global crop production. *J. Clean. Prod.* **2014**, *73*, 52–62. [CrossRef]
19. Berger, M.; Finkbeiner, M. Methodological challenges in volumetric and impact-oriented water footprints. *J. Ind. Ecol.* **2013**, *17*, 79–89. [CrossRef]
20. Hoekstra, A.Y. A critique on the water-scarcity weighted water footprint in LCA. *Ecol. Indic.* **2016**, *66*, 564–573. [CrossRef]

21. Pfister, S.; Hellweg, S. *Surface Water Use-Human Health Impacts*; LC-IMPACT Project(EC:FP7); LC-IMPACT: Zürich, Switzerland, 2011.
22. Nunez, M.; Pfister, S.; Vargas, M.; Anton, A. Spatial and temporal specific characterisation factors for water use impact assessment in Spain. *Int. J. LCA* **2015**, *20*, 128–138. [CrossRef]
23. Scherer, L.; Pfister, S. Dealing with uncertainty in water scarcity footprints. *Environ. Res. Lett.* **2016**, *11*, 054008. [CrossRef]
24. Chen, X.; Corson, M.S. Influence of emission-factor uncertainty and farm-characteristic variability in LCA estimates of environmental impacts of french dairy farms. *J. Clean. Prod.* **2014**, *81*, 150–157. [CrossRef]
25. Groen, E.A.; Heijungs, R.; Bokkers, E.A.M.; de Boer, I.J.M. Methods for uncertainty propagation in life cycle assessment. *Environ. Model. Softw.* **2014**, *62*, 316–325. [CrossRef]
26. Ono, Y. Water Inventory Development for Application of Water Footprint. Bachelor's Thesis, Tokyo City University, Yokohama, Japan, 2010.
27. Rural Research Institute(RRI). *Water Footprint Estimation and Application for Sustainable Water Resources Use*; Rural Research Institute(RRI): Ansan, Korea, 2014. (In Korean)
28. Water Resource Management Information System. Available online: http://wamis.go.kr/eng/ (accessed on 16 March 2017).
29. Rural Agricultural Water Resource Information System. Available online: http://rawris.ekr.or.kr/ (accessed on 16 March 2017).
30. Water Use in Life Cycle Assessment(WULCA). Available online: http://www.wulca-waterlca.org/aware.html (accessed on 26 March 2017).
31. Jung, I.W.; Bae, D.H. A study on PRMS applicability for Korean River Basin. *J. Korea Water Resour. Assoc.* **2005**, *38*, 713–725. (In Korean) [CrossRef]
32. OECD. *Water Resources Allocation: Sharing Risks and Opportunities*; OECD Publishing: Paris, France, 2015. [CrossRef]
33. The Ministry of Land Transport and Maritime Affairs. *Comprehensive Water Resource Plan-Water Vision 2020*; The Ministry of Land Transport and Maritime Affairs: Sejong City, Korea, 2016. (In Korean)
34. Pastor, A.V.; Ludwig, F.; Biemans, H.; Hoff, H.; Kabat, P. Accounting for environmental flow requirements in global water assessments. *Hydrol. Earth Syst. Sci.* **2013**, *18*, 5041–5059. [CrossRef]
35. Martorell, S.; Guedes Soares, C.; Barnett, J. *Safety, Reliability and Risk Analysis*, 1st ed.; CRC Press: London, UK, 2009; p. 2951, ISBN 978-0-415-48792-4.
36. Bevington, P.R.; Robinson, D.K. *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed.; McGraw-Hill: New York, NY, USA, 2003; pp. 75–97, ISBN 978-0-07-247227-1.
37. Vořechovský, M.; Novák, D. Statistical correlation in stratified sampling. In Proceedings of the International Conference on Applications of Statistics and Probability in Civil Engineering, San Francisco, CA, USA, 6–9 July 2003; MillPress: Rotterdam, The Netherlands, 2003; pp. 119–124.
38. Vidyamurthy, G. *Pairs Trading: Quantitative Methods and Analysis*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 2004; pp. 37–51, ISBN 978-0471460671.
39. Anderson, T.W.; Darling, D.A. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Ann. Math. Stat.* **1952**, *23*, 193–212. [CrossRef]
40. Intergovernmental Panel ON Climate Change (IPCC). *Guidelines for National Greenhouse Gas Inventories Chapter 3 of Volume 1*; Intergovernmental Panel ON Climate Change (IPCC): Geneva, Switzerland, 2006.
41. Jolliet, O.; Saade-Sbeih, M.; Shaked, S.; Jolliet, A.; Crettaz, P. *Environmental Life Cycle Assessment*, 1st ed.; CRC Press: New York, NY, USA, 2015; pp. 170–179, ISBN 978-1-4398-8770-7.
42. Pattengale, N.D.; Alipour, M.; Bininda-Emonds, O.R.P.; Moret, B.M.E.; Stamatakis, A. How Many Bootstrap Relicates Are Necessary? In *Proceedings of the Computatinal Molecular Biology*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 184–200.
43. Orloff, J.; Bloom, J. *18.05 Introduction to Probability and Statistics*; Massachusetts Institute of Technology, MIT OpenCourseWare: Massachusetts, MA, USA, 2014.
44. Lahiri, S.N. *Resampling Methods for Dependent Data*, 1st ed.; Springer: New York, NY, USA, 2003; pp. 1–21, ISBN 978-1-4419-1848-2.
45. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Application*, 1st ed.; CAMBRIDGE University Press: New York, NY, USA, 1997; pp. 385–436, ISBN 978-0-521-57471-6.
46. Oracle Crystal Ball. *Classroom Edition 11.1.4100.0*; Oracle Corporation: Redwood City, CA, USA, 2014.

47. Lee, M.H.; Lee, J.S.; Lee, J.Y.; Kim, Y.H.; Park, Y.S.; Lee, K.M. Uncertainty analysis of a GHG emission model output using the block bootstrap and monte carlo simulation. *Sustainability* **2017**, *9*, 1522. [CrossRef]
48. Gramacki, A. *Nonparametric Kernel Density Estimation and Its Computational Aspects*, 1st ed.; Springer: New York, NY, USA, 2018; pp. 7–24, ISBN 978-3-319-71688-6.
49. Lloyd, S.M.; Ries, R. Characterizing, propagating, and analyzing uncertainty in life-cycle assessment. *J. Ind. Ecol.* **2007**, *11*, 161–179. [CrossRef]
50. Bertsekas, D.P.; Tsitsiklis, J.N. *Introduction to Probability*, 2nd ed.; Athena Scientific: Belmont, MA, USA, 2008; pp. 263–293, ISBN 978-1-886529-38-0.